

RECEIVE

MAY 16 2002

TECH CENTER 1600/2900



1600

ENTERED

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/706,243A

DATE: 05/09/2002  
TIME: 13:49:09

Input Set : A:\seqlist.txt  
Output Set: N:\CRF3\05092002\I706243A.raw

3 <110> APPLICANT: Cox III, George N.  
4 Case, Casey Christopher  
5 Eisenberg, Stephen P.  
6 Jarvis, Eric E.  
7 Spratt, Sharon K.  
9 <120> TITLE OF INVENTION: REGULATION OF ENDOGENOUS GENE EXPRESSION IN CELLS  
10 USING ZINC FINGER PROTEINS  
12 <130> FILE REFERENCE: 8325-0002.10 / S2-US3  
14 <140> CURRENT APPLICATION NUMBER: 09/706,243A  
15 <141> CURRENT FILING DATE: 2000-11-03  
17 <160> NUMBER OF SEQ ID NOS: 43  
19 <170> SOFTWARE: PatentIn Ver. 2.0  
21 <210> SEQ ID NO: 1  
22 <211> LENGTH: 25  
23 <212> TYPE: PRT  
24 <213> ORGANISM: Artificial Sequence  
26 <220> FEATURE:  
27 <223> OTHER INFORMATION: Description of Artificial Sequence:exemplary motif  
28 of C2H2 class of zinc finger proteins (ZFP)  
30 <220> FEATURE:  
31 <221> NAME/KEY: MOD\_RES  
32 <222> LOCATION: (2)..(3)  
33 <223> OTHER INFORMATION: Xaa = any amino acid  
35 <220> FEATURE:  
36 <221> NAME/KEY: MOD\_RES  
37 <222> LOCATION: (4)..(5)  
38 <223> OTHER INFORMATION: Xaa = any amino acid, may be present or absent  
40 <220> FEATURE:  
41 <221> NAME/KEY: MOD\_RES  
42 <222> LOCATION: (7)..(18)  
43 <223> OTHER INFORMATION: Xaa = any amino acid  
45 <220> FEATURE:  
46 <221> NAME/KEY: MOD\_RES  
47 <222> LOCATION: (20)..(22)  
48 <223> OTHER INFORMATION: Xaa = any amino acid  
50 <220> FEATURE:  
51 <221> NAME/KEY: MOD\_RES  
52 <222> LOCATION: (23)..(24)  
53 <223> OTHER INFORMATION: Xaa = any amino acid, may be present or absent  
55 <400> SEQUENCE: 1  
W--> 56 Cys Xaa Xaa Xaa Xaa Cys Xaa  
57 1 5 10 15  
W--> 59 Xaa Xaa His Xaa Xaa Xaa Xaa Xaa His

RAW SEQUENCE LISTING DATE: 05/09/2002  
PATENT APPLICATION: US/09/706,243A TIME: 13:49:09

Input Set : A:\seqlist.txt  
Output Set: N:\CRF3\05092002\I706243A.raw

60 20 25  
61 <210> SEQ ID NO: 2  
62 <211> LENGTH: 10  
63 <212> TYPE: DNA  
64 <213> ORGANISM: Artificial Sequence  
66 <220> FEATURE:  
67 <223> OTHER INFORMATION: Description of Artificial Sequence:ZFP target site  
68 with two overlapping D-able subsites  
70 <220> FEATURE:  
71 <221> NAME/KEY: modified\_base  
72 <222> LOCATION: (1)..(2)  
73 <223> OTHER INFORMATION: n = g,a,c or t  
75 <220> FEATURE:  
76 <221> NAME/KEY: modified\_base  
77 <222> LOCATION: (5)  
78 <223> OTHER INFORMATION: n = g,a,c or t  
80 <220> FEATURE:  
81 <221> NAME/KEY: modified\_base  
82 <222> LOCATION: (8)  
83 <223> OTHER INFORMATION: n = g,a,c or t  
85 <220> FEATURE:  
86 <221> NAME/KEY: modified\_base  
87 <222> LOCATION: (9)  
88 <223> OTHER INFORMATION: n = a,c or t; if g, then position 10 cannot be g  
89 or t  
91 <220> FEATURE:  
92 <221> NAME/KEY: modified\_base  
93 <222> LOCATION: (10)  
94 <223> OTHER INFORMATION: n = a or c; if g or t, then position 9 cannot be g  
96 <400> SEQUENCE: 2  
W--> 97 nngkngknnn 10  
100 <210> SEQ ID NO: 3  
101 <211> LENGTH: 10  
102 <212> TYPE: DNA  
103 <213> ORGANISM: Artificial Sequence  
105 <220> FEATURE:  
106 <223> OTHER INFORMATION: Description of Artificial Sequence:ZFP target site  
107 with three overlapping D-able subsites  
109 <220> FEATURE:  
110 <221> NAME/KEY: modified\_base  
111 <222> LOCATION: (1)..(2)  
112 <223> OTHER INFORMATION: n = g,a,c or t  
114 <220> FEATURE:  
115 <221> NAME/KEY: modified\_base  
116 <222> LOCATION: (5)  
117 <223> OTHER INFORMATION: n = g,a,c or t  
120 <220> FEATURE:  
121 <221> NAME/KEY: modified\_base  
122 <222> LOCATION: (8)

RAW SEQUENCE LISTING DATE: 05/09/2002  
PATENT APPLICATION: US/09/706,243A TIME: 13:49:09

Input Set : A:\seqlist.txt  
Output Set: N:\CRF3\05092002\I706243A.raw

123 <223> OTHER INFORMATION: n = g,a,c or t  
125 <400> SEQUENCE: 3  
W--> 126 nngkngkngk 10  
129 <210> SEQ ID NO: 4  
130 <211> LENGTH: 5  
131 <212> TYPE: PRT  
132 <213> ORGANISM: Artificial Sequence  
134 <220> FEATURE:  
135 <223> OTHER INFORMATION: Description of Artificial Sequence:linker  
137 <400> SEQUENCE: 4  
138 Asp Gly Gly Gly Ser  
139 1 5  
142 <210> SEQ ID NO: 5  
143 <211> LENGTH: 5  
144 <212> TYPE: PRT  
145 <213> ORGANISM: Artificial Sequence  
147 <220> FEATURE:  
148 <223> OTHER INFORMATION: Description of Artificial Sequence:linker  
150 <400> SEQUENCE: 5  
151 Thr Gly Glu Lys Pro  
152 1 5  
155 <210> SEQ ID NO: 6  
156 <211> LENGTH: 9  
157 <212> TYPE: PRT  
158 <213> ORGANISM: Artificial Sequence  
160 <220> FEATURE:  
161 <223> OTHER INFORMATION: Description of Artificial Sequence:linker  
163 <400> SEQUENCE: 6  
164 Leu Arg Gln Lys Asp Gly Glu Arg Pro  
165 1 5  
168 <210> SEQ ID NO: 7  
169 <211> LENGTH: 4  
170 <212> TYPE: PRT  
171 <213> ORGANISM: Artificial Sequence  
173 <220> FEATURE:  
174 <223> OTHER INFORMATION: Description of Artificial Sequence:linker  
176 <400> SEQUENCE: 7  
177 Gly Gly Arg Arg  
178 1  
181 <210> SEQ ID NO: 8  
182 <211> LENGTH: 5  
183 <212> TYPE: PRT  
184 <213> ORGANISM: Artificial Sequence  
186 <220> FEATURE:  
187 <223> OTHER INFORMATION: Description of Artificial Sequence:linker  
189 <400> SEQUENCE: 8  
190 Gly Gly Gly Gly Ser  
191 1 5  
194 <210> SEQ ID NO: 9

RAW SEQUENCE LISTING DATE: 05/09/2002  
PATENT APPLICATION: US/09/706,243A TIME: 13:49:09

Input Set : A:\seqlist.txt  
Output Set: N:\CRF3\05092002\I706243A.raw

195 <211> LENGTH: 8  
196 <212> TYPE: PRT  
197 <213> ORGANISM: Artificial Sequence  
199 <220> FEATURE:  
200 <223> OTHER INFORMATION: Description of Artificial Sequence:linker  
202 <400> SEQUENCE: 9  
203 Gly Gly Arg Arg Gly Gly Ser  
204 1 5  
207 <210> SEQ ID NO: 10  
208 <211> LENGTH: 9  
209 <212> TYPE: PRT  
210 <213> ORGANISM: Artificial Sequence  
212 <220> FEATURE:  
213 <223> OTHER INFORMATION: Description of Artificial Sequence:linker  
215 <400> SEQUENCE: 10  
216 Leu Arg Gln Arg Asp Gly Glu Arg Pro  
217 1 5  
220 <210> SEQ ID NO: 11  
221 <211> LENGTH: 12  
222 <212> TYPE: PRT  
223 <213> ORGANISM: Artificial Sequence  
225 <220> FEATURE:  
226 <223> OTHER INFORMATION: Description of Artificial Sequence:linker  
228 <400> SEQUENCE: 11  
229 Leu Arg Gln Lys Asp Gly Gly Ser Glu Arg Pro  
230 1 5 10  
233 <210> SEQ ID NO: 12  
234 <211> LENGTH: 16  
235 <212> TYPE: PRT  
236 <213> ORGANISM: Artificial Sequence  
238 <220> FEATURE:  
239 <223> OTHER INFORMATION: Description of Artificial Sequence:linker  
242 <400> SEQUENCE: 12  
243 Leu Arg Gln Lys Asp Gly Gly Ser Gly Gly Ser Glu Arg Pro  
244 1 5 10 15  
247 <210> SEQ ID NO: 13  
248 <211> LENGTH: 25  
249 <212> TYPE: DNA  
250 <213> ORGANISM: Artificial Sequence  
252 <220> FEATURE:  
253 <223> OTHER INFORMATION: Description of Artificial Sequence:ZFP target site  
254 region surrounding initiation site of vascular  
255 endothelial growth factor (VEGF) gene containing  
256 two 9-base pair target sites  
258 <220> FEATURE:  
259 <221> NAME/KEY: protein\_bind  
260 <222> LOCATION: (4)..(12)  
261 <223> OTHER INFORMATION: upstream 9-base pair ZFP VEGF1 target site  
263 <220> FEATURE:

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/706,243A

DATE: 05/09/2002

TIME: 13:49:09

Input Set : A:\seqlist.txt

Output Set: N:\CRF3\05092002\I706243A.raw

264 <221> NAME/KEY: protein\_bind  
 265 <222> LOCATION: (14)..(22)  
 266 <223> OTHER INFORMATION: downstream 9-base pair ZFP VEGF3a target site  
 268 <400> SEQUENCE: 13  
 269 agcgccccagg atcgccggagg cttgg 25  
 272 <210> SEQ ID NO: 14  
 273 <211> LENGTH: 298  
 274 <212> TYPE: DNA  
 275 <213> ORGANISM: Artificial Sequence  
 277 <220> FEATURE:  
 278 <223> OTHER INFORMATION: Description of Artificial Sequence:VEGF1 ZFP  
 279 construct targeting upstream 9-base pair target  
 280 site in VEGF promoter  
 282 <220> FEATURE:  
 283 <221> NAME/KEY: CDS  
 284 <222> LOCATION: (2)..(298)  
 285 <223> OTHER INFORMATION: VEGF1  
 287 <400> SEQUENCE: 14  
 288 g gta ccc ata cct ggc aag aag aag cag cac atc tgc cac atc cag ggc 49  
 289 Val Pro Ile Pro Gly Lys Lys Gln His Ile Cys His Ile Gln Gly  
 290 1 5 10 15  
 292 tgt ggt aaa gtt tac ggc aca acc tca aat ctg cgt cgt cac ctg cgc 97  
 293 Cys Gly Lys Val Tyr Gly Thr Ser Asn Leu Arg Arg His Leu Arg  
 294 20 25 30  
 296 tgg cac acc ggc gag agg cct ttc atg tgt acc tgg tcc tac tgt ggt 145  
 297 Trp His Thr Gly Glu Arg Pro Phe Met Cys Thr Trp Ser Tyr Cys Gly  
 298 35 40 45  
 300 aaa cgc ttc acc cgt tcg tca aac ctg cag cgt cac aag cgt acc cac 193  
 301 Lys Arg Phe Thr Arg Ser Ser Asn Leu Gln Arg His Lys Arg Thr His  
 302 50 55 60  
 303 acc ggt gag aag aaa ttt gct tgc ccg gag tgt ccg aag cgc ttc atg 241  
 304 Thr Gly Glu Lys Lys Phe Ala Cys Pro Glu Cys Pro Lys Arg Phe Met  
 305 65 70 75 80  
 307 cgt agt gac cac ctg tcc cgt cac atc aag acc cac cag aat aag aag 289  
 308 Arg Ser Asp His Leu Ser Arg His Ile Lys Thr His Gln Asn Lys Lys  
 309 85 90 95  
 311 ggt gga tcc 298  
 312 Gly Gly Ser  
 315 <210> SEQ ID NO: 15  
 316 <211> LENGTH: 99  
 317 <212> TYPE: PRT  
 318 <213> ORGANISM: Artificial Sequence  
 320 <220> FEATURE:  
 321 <223> OTHER INFORMATION: Description of Artificial Sequence:VEGF1 ZFP  
 322 construct targeting upstream 9-base pair target  
 323 site in VEGF promoter  
 325 <400> SEQUENCE: 15  
 326 Val Pro Ile Pro Gly Lys Lys Gln His Ile Cys His Ile Gln Gly  
 327 1 5 10 15

RAW SEQUENCE LISTING ERROR SUMMARY                    DATE: 05/09/2002  
PATENT APPLICATION: US/09/706,243A                    TIME: 13:49:10

Input Set : A:\seqlist.txt  
Output Set: N:\CRF3\05092002\I706243A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:1; Xaa Pos. 2,3,4,5,7,8,9,10,11,12,13,14,15,16,17,18,20,21,22,23,24  
Seq#:2; N Pos. 1,2,5,8,9,10  
Seq#:3; N Pos. 1,2,5,8  
Seq#:41; N Pos. 1,2  
Seq#:42; N Pos. 1,2  
Seq#:43; N Pos. 1,2

**VARIABLE LOCATION, SUMMARY** DATE: 05/09/2002  
**PATENT APPLICATION:** US/09/706,243A TIME: 13:49:10

Input Set : A:\seqlist.txt  
Output Set: N:\CRF3\05092002\I706243A.raw

### Use of n's or Xaa's (NEW RULES):

Use of n's and/or Xaa's have been detected in the Sequence Listing.

Use of <220> to <223> is MANDATORY if n's or Xaa's are present.

in <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.

Seq#:1; Xaa Pos. 2,3,4,5,7,8,9,10,11,12,13,14,15,16,17,18,20,21,22,23,24

Seq#:2; N Pos. 1,2,5,8,9,10

Seq#:2: N Pos. 1,2,5,8

Seq#:3: N Pos. 1,2;

Seq#:41, N POS. 1,2

Seq#:42; N POS. 1,2

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/09/706,243A

DATE: 05/09/2002

TIME: 13:49:10

Input Set : A:\seqlist.txt

Output Set: N:\CRF3\05092002\I706243A.raw

L:56 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:0  
L:59 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:16  
L:97 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:0  
L:126 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:0  
L:861 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:41  
L:861 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:41  
L:861 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:41 after pos.:0  
L:873 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:42  
L:873 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:42  
L:873 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:42 after pos.:0  
L:885 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:43  
L:885 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:43  
L:885 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:43 after pos.:0